

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: August 23, 2003, 18:20:39 ; Search time 1902 Seconds
(without alignments)
4472.433 Million cell updates/sec

Title: US-09-745-506-37
Perfect score: 1799
Sequence: 1 MDKALLSLNDFASLSPAE.....LENNKINIIILSETRDPLQV 350

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlh
-Q/cgn2.1/USPRO.spool/US09745506/runat.22082003.104401.7049/app-query.fasta.1.519
-DB-EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR.SCORE=dot -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSITE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09745506 -GCCN.1.1.2135 -runat.22082003.104401.7049 -NCPD=6 -ICPD=3
-NO_MAP -LARGEDBERR -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGCLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_estbun:*
3: em_estcin:*
4: em_estcnu:*
5: em_estcov:*
6: em_estcpl:*
7: em_estcro:*
8: em_hc1:*
9: gb_estc1:*
10: gb_estc2:*
11: gb_hc1:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estcin:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_luv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1599	88.9	1201	9	AL529615
2	1548	86.0	1376	11	AK011670
3	1539	85.5	1117	12	BM545164
4	1534	85.3	1201	13	BX446370
5	1481	82.3	1201	9	AL581997
6	1407	78.2	1467	11	AK003978
7	1378	76.6	930	13	BX393871
8	1363	75.8	857	13	BUI72435
9	1362	75.7	929	13	BUI859307
10	1362	75.7	985	13	BQ927763
11	1360.5	75.6	836	13	BQ229243
12	1345.5	74.8	1201	9	AL582089
13	1335	74.2	1201	9	AL563496
14	1316	73.2	1919	11	AK076788
15	1315	73.1	957	9	AL521920
16	1310	72.8	845	10	BG754550
17	1298	72.2	1201	13	BX377118
18	1258	69.9	742	12	B1093955
19	1251.5	69.6	1201	9	AL523985
20	1233.5	68.6	1019	12	BM557530
21	1233	68.5	890	9	AL520538
22	1201	66.8	888	13	BUI95469
23	1198	66.6	732	10	BG472953
24	1197	66.5	980	13	BQ854847
25	1194	66.4	905	10	BF973755
26	1192.5	66.3	843	12	BI753964
27	1182	65.7	958	10	BF973889
28	1177	65.4	817	12	BI755123
29	1172	65.1	977	10	BF663320
30	1162.5	64.6	1079	13	BX339175
31	1158.5	64.4	963	13	BE797115
32	1157	64.3	759	10	BE275324
33	1135.5	63.1	977	14	CD513847
34	1132.5	63.0	843	13	BQ848893
35	1126.5	62.6	909	14	CD385139
36	1118	62.1	961	9	AL521919
37	1114	61.9	984	13	BX385101
38	1103	61.3	909	13	BUI65031
39	1102	61.3	999	10	BE745082
40	1097	61.0	946	13	BQ963633
41	1082.5	60.2	945	10	BE747311
42	1077	59.9	772	14	CA321994
43	1075	59.8	877	13	BQ437698
44	1069.5	59.4	768	12	BI772801
45	1062	59.0	705	10	BE745354

ALIGNMENTS

RESULT 1
AL529615
LOCUS 1201 bp mRNA linear EST 23-MAY-2003
DEFINITION AL529615 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
CDNA clone CS0DD006YH18 5-PRIME, mRNA sequence.
ACCESSION AL529615
VERSION AL529615
KEYWORDS GI:31067458
SOURCE EST.
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1201)

AUTHORS Li, M.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT On Feb 13, 2001 this sequence version replaced g1:12793308

Genoscope - Centre National de Séquençage
Contact: Genoscope
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 1287.f For more information about this cluster, see

cg1-bin/cluster_cg1seq-CS0DD006DD09GP1.kcluster-1287.f. contact :
Feng Liang Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com> Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DD006DD09GP1.
T000510070001.f.1287

FEATURES	Location/Qualifiers
source	1. .1201

```

/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="CSDD0006yh18"
/tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"
/clone_1ib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-0190(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      319 a      286 c      274 g      299 t      23 others
ORIGIN

```

Alignment Scores:	
Pred. No.:	3,53e-162
Score:	1599.00
Percent Similarity:	96.658
Best Local Similarity:	96.048
Query Match:	88.888
DB:	9
Length:	120
Matches:	315
Conservative:	2
Mismatches:	11
Indels:	0
Gaps:	0

US-09-745-506-37 (1-350) x AL529615 (1-1201)

QY	1	MetIAspLeuValAlaLeuLeuSerSerLeuAsnAspPheAlaSerLeuSerPheAlaLeu	20
Db	171	ATGGAATTGGAAGCTCTCCCTTTCTCTTGAAAGACTTTGGACACCCCTCGTTGGCTAG	233
QY	21	SetrTrpAspAsnAlaGlyLeuLeuValGluProSerProProHisIsthValAsnThrIleu	40
Db	231	AGTAGGACCAATGTTGGATTACTGGTGGAAACCAAGCCCAACATATCTGAATACATCTC	299
QY	41	PheLeuThrAsnAspLeuThrGluGluValMetGluGluValLeuGlnIysIysAlaAsp	60
Db	291	TTCTGTGCACATACCTCGACTGAGGAAGAGTAGAGAGAGTCTCAAAAGAGGCACAC	355
QY	61	LeuIleLeuSerThrIHisProProIlePheArgProMetIysArgIleThrTrpAsnThr	80
Db	351	CTCATTTCTCTCCATCCATCCGCGCTACTTCCGACCATATGAACGCATACCTGGAAACA	411
QY	81	TrpIysGluArgLeuValIleArgAlaLeuGluAsnArgValGlyIleTyrSerProHis	100
Db	411	TGGAAAGAGGCGCTGGTGATCCGCGCTCGAGAAACAGATCGTATCTACTCTCCAT	477
QY	101	ThrIleAlaTrpAspAlaIleProGlnIysValAsnAspTrpLeuAlaIysGlyLeuGlnIysAla	120
Db	471	ACACACCTATATGCTGGGCCCGCAGGCGCTCAACATCGTTGGCTTAAAGGGCTTGAGCT	533
QY	121	CysThrSerArgProIleHisIProSerIysAlaIleProAsnTyrProThrGluIysAsnHis	140
Db	531	TGTACCTCCAGGCGCCATACATCTCTCCAAAGCTCCCAACTACCTCAAGAGGAACAC	599
QY	141	ArgValGluIueHisAsnValAsnTyrThrGlnAspLeuAspLysValMetSerAlaValIys	160
Db	591	CGAATTAAGATTCACGTTAATCAACCCAGACCTGGACAAGAATGATGCTGCACTGTGAA	655

Oy		161	GlylaesplgylaservatlrSerPhaserlargrhrglansgluglnrthr	180
Db		651	GGAAATTACCGGTGTTTCTGCACCTCTTTTCTCCTAGACTGGTAATGAGAACAAACA	710
Oy		181	ArglleasnleuaancysrrhnglnLysalaleumetglnvalValaaSpheLeuSerarg	200
Db		711	CGGAATTATCATCGAATTGTACTCAAGAAGCGTTTGATCGAGGTGTAAGATTTTCCTCCGG	770
Oy		201	AsnLysglnleuTyrglnLysThrhglnlleuSerLeuGlulysProleuLeuHis	220
Db		771	AACAACAACATTVAATFCAGAAAGACGAAATTCCTGCACCTGGAGAAAGCCCTTGCTCTACAT	830
Oy		221	ThrslymetgLyargleucysthrLeuaaspLuservalSerleualathrmelileasp	240
Db		831	ACTGGATGGGACGGTATGACACACTGGATGAATCTGTCTCCCGCACCATATMNGAT	890
Oy		241	ArglleasArghlsleuLysLeuSerHisIleArgleualaLeuGlyValArgYthr	260
Db		891	CGAATTAATAAACACACCTAAACATCTCATATTCGTAGGCCCTTGCGGTGGAGAAC	950
Oy		261	LeucluserGlnvalLysValAlaAleucysAlaglyserGlyserSerValLeuGln	280
Db		951	TTAAGATCTCAAGCAACAAGTGTGGCCCTGTGTCCGGTCTGGAGACAGCTTCGCAG	1010
Oy		281	GlyValgluaAlaaspLeuTyrlenThrGlylglnMetSerHisIshASpThrLeuaSPala	300
Db		1011	GGTGTGGAGCGTGCACCTTTACCCTCAGAGTAGAMGTCCCMCAMGAATTTTGGATGCT	1070
Oy		301	AlaseRnglnglylleasnValilleuencysgluhIsSerAaThrhgUArGlyPheLeu	320
Db		1071	GCTTCCCACARATVAATGATGTAATCTCTGTGAAMACAGCAACACTGAAGAGSGSTKTYTT	1130
Oy		321	SerAspleuArgaSpmleuASP 328	
Db		1131	TCTGACCTTCGAGAAAMCGTGGAT 1154	
LOCUS	AKO11670	1376 bp	mRNA	linear HTC 05-DEC-2002
DEFINITION	Mus musculus 10 days embryo whole body cDNA, RIKEN full-length enriched library, clone:2610034L10 product:Nggl interacting factor 3-like 1 (S. pombe), full insert sequence.			
ACCESSION	AKO11670			
VERSION	AKO11670.1 GI:12847943			
KEYWORDS	HTC; CAP trapper.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	Carninci,P. and Hayashizaki,Y.			
TITLE	High-efficiency full-length cDNA cloning			
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)			
PUBMED	99279253			
REFERENCE	10349636			
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.			
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes			
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)			
MEDLINE	20499374			
PUBMED	11042159			
REFERENCE	3			
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Kono,H., Akiyama,J., Nishi,K., Kitsuna,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kasaiwagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Isikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kiru,A. and Hayashizaki,Y.			
Riken integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer				

||||:|||||
Db 606 GGGGTGGAAGTGTCTGTCTACTCTTCTTCTCCAGGTGTGATGGTGAAGCAACACC 665
Qy 181 Arg1leasnleuasnCysThrGlnLysAlaLeuMetGlnValAlaAspPheLeuSerArg 200
Db 666 CGGATTCAGCCGGAATGTGTACTACACAGACTTTATGCAAGTGCAGAGCTTTCTTCCAG 725
Qy 201 AsnLysGlnLeuTyrgLysThrGlnLysLeuSerLeuGlnLysProLeuLeuHis 220
Db 726 GACAGACAACTTATTCAGAAATGAAATCTTTCATTGGAAGACCTTCTTTCAT 785
Qy 221 ThrGlnMetGlyArgLeuCysThrLeuAspLysSerValSerLeuAlaThrMetLeasp 240
Db 786 ACTGGAATGGAGCGGTGTGTGCACACTGATGAATCTGTCTCCGCAATATATGATAG 845
Qy 241 Arg1leTyArgH1sLeuLysLeuSerHis1leArGlnLeuAlaLeuGlyValArgThr 260
Db 846 CGAATCAAAAACACACCTTAAGCTGTGCAATCTTCCGCTTACGCTTGGAGTGGAGACA 905
Qy 261 LeuGlnSerGlnValLysValAlaLeuCysAlaGlySerGlySerValLeuGln 280
Db 906 TTAGAGTCCCAAGCAAGTGTGTGCGCTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGT 965
Qy 281 GlyValGlnAlaAspLeuTyrgLysThrGlnLysLeuSerHis1AspPheLeuAspLys 300
Db 966 GGAATGGAGCGCGACCTTACCTACACAGTGAATGTCCACCAATGATGTCTGTGATGCT 1025
Qy 301 AlaSerGlnGlyLeuAsnVal1leLeuGlnLysSerAsnThrGlnArgGlyPheLeu 320
Db 1026 GCTTCCAAAGGATCAATGATCTATCTTGTGACACACACACACACACACACACACAC 1085
Qy 321 SerAspLeuArgAspMetLeuAspSerHisLeuGlnAsnLys1leAsn1leLeuSer 340
Db 1086 TCTGAGCTTCAAGAAATCTGGGTTCACCTTGTGATTAAGATTAACTATCTGTCT 1145
Qy 341 GlnThrAspArgAspProLeuGlnValVal 350
Db 1146 GAGACAGACAGGAGCCCTCTCGTGTGTGT 1175

RESULT 3
BM545164 1117 bp mRNA linear EST 20-FeB-2002
LOCUS AGENCOURT.6497454 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5588980
DEFINITION AGNCOURT.6497454 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5588980
ACCESSION BM545164
VERSION BM545164.1 GI:18777026
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1117)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: csapbs-remail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12360 row: P column: 05
High quality sequence stop: 734.

FEATURES
Source Location/Qualifiers
1..1117
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5588980"
/lab_host="DH10B"

/clone_11b="NIH_MGC_125"
/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
Site_1: EcoRV (destroyed); Site_2: NotI; RNA source pool
of three ovaries, from females ranging in age from 38 to
49 yo. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 2.1 kb, insert size range 1-3.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 036."

BASE COUNT 278 a 282 c 260 g 293 t 4 others
ORIGIN
Alignment Scores:
Pred. No.: 9.82e-156 Length: 1117
Score: 1539.00 Matches: 312
Percent Similarity: 95.43% Conservative: 1
Best Local Similarity: 95.12% Mismatches: 10
Query Match: 85.55% Indels: 5
Gaps: 0
US-09-745-506-37 (1-350) x BM545164 (1-1117)

Qy 1 MetAspLeuLysAlaLeuLeuSerSerLeuAsnAspPheAlaSerLeuSerPheAlaGln 20
Db 88 ATGGATTTGAAGGCTCTCTCTTCTCTTCTGTAATGACTTTGCAATCCCTCTGCTGTGAG 147
Qy 21 SerTrpAspAsnValGlyLeuLeuValGlnProSerProProHisThrValAsnThrLeu 40
Db 148 AGTTGGGCAATGTTGGATTACTGTGTGAGAACCAAGCCACCAACATGTAATATACACTC 207
Qy 41 PheLeuThrAsnAspLeuThrGlnGlnValMetGlnValAlaLeuGlnLysLysAlaAsp 60
Db 208 TTCCGTGACCAATGACCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 267
Qy 61 Leu1leLeuSerTyrgH1sProPro1lePheArgProMetLysArgGlyLeuThrAsnThr 80
Db 268 CTCATCTCTCTTACACATCGCGCTTATCTTCCGACCAAGAGCGATTAACCTGGACACA 327
Qy 81 TrpLysGlnArgLeuVal1leArgAlaLeuGlnAsnArgValGlyLeuSerProHis 100
Db 328 TGGAGAGAGCGCGCTGTGTATCCGGGCTGTGGAGAACAGACGATTAATCTCTCTCAT 387
Qy 101 ThrAlaTyrgAspAlaAlaProGlnGlyValAsnAsnTrpLeuAlaLysGlyLeuGlyAla 120
Db 388 ACAGCCATGATGCTGCGCCCGCAGGCGCTCAACAACATGTTGGCTAAAGGCGCTTGACCT 447
Qy 121 CysThrSerArgPro1leHisProSerTyrgAlaProAsnTyrgProThrGlnGlyAsnHis 140
Db 448 TGTACCTCCAGGCCCATCATCTCTTCCAAAGCTCCCACTACCTACAGAGGAGAACAC 507
Qy 141 ArgValGlnPheAsnValAsnTyrgThrGlnAspLeuAspLysValMetSerAlaValLys 160
Db 508 CGAGTAGAATTCACAGCTTAACCTACACCCAGACCGTGGACCAAGCATGTCTGCACTGAA 567
Qy 161 Gly1leAspGlyValSerValThrSerPheSerAlaArgThrGlyAsnGlnGlnThr 180
Db 568 GGAATTCAGCGGTGTCTGTCTCTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 627
Qy 181 Arg1leAsnLeuAsnCysThrGlnLysAlaLeuMetGlnValAlaAspPheLeuSerArg 200
Db 628 CGGATTAATCTGAAATGTGTACTACAGAGCTTTGATGCAAGTGTGATTTCTTCCCGG 687
Qy 201 AsnLysGlnLeuTyrgLysThrGlnLysLeuSerLeuGlnLysProLeuLeuHis 220
Db 688 AACCAACAACTTATTCAGAGACGGAATTCGTGCACTGGAAGACCTTCTCTCTACAT 747
Qy 221 ThrGlnMetGlyArgLeuCysThrLeuAspLysSerValSerLeuAlaThrMetLeasp 240
Db 748 ACTGGAATGGAGCGGTGTGTGCACTGATGATGTCTCTCCGCGACACATGATATAT 807
Qy 241 Arg1leTyrgH1sLeuLysLeuSerHis1leArGlnLeuAlaLeuGlyValArgThr 260

Db 808 CGAATANAGACACCTAAACATCATATTCGCTAGCCCTTTGGGGTGGGAGAAC 867
QY 260 rleucguserglnvallyvalaala-leucysalaiglysergyservallleug 280
Db 868 CTTAAGTCTCAAGCTCAAGTGTGGCCCTGTGTCTGTCTGTGGACAGCTTCTGC 927
QY 280 lgllyvalgluialaspleuTYrleuThglglumetserhshisapThr-leuasP 299
Db 928 AGGTGTTCAGCTGACCTTACTTACACAGTGAGATGCCATCATCTGATCTTTGGAT 987
QY 300 AlaAlaSerGlnGlylleasn-valilleucysglnhissersasn-ThrgluarGlyP 319
Db 988 GCTGCTNCCCAAGATNATGCGCATCTCTGTGAAACACAGCAACCTGGAACGAGCT 1047
QY 319 heluseraspleu 323
Db 1048 TTCTTCTGACCTT 1061

RESULT 4

LOCUS BX446370 1201 bp mRNA linear EST 22-MAY-2003
DEFINITION BX446370 Homo sapiens PLACENTA Homo sapiens CDNA clone CL0BA0042B11
5-PRIME, mRNA sequence.

ACCESSION BX446370
VERSION BX446370.1 GI:31023721
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Li, M.B., Gruber, C., Jesssee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished

COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1287.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgl-bin/cluster.cgi?seq=CL0BA0042B11RPL&cluster=1287.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CL0BA0042B11RPL.

FEATURES

source

1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CL0BA0042B11"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: PCMVSPORT_6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

BASE COUNT 349 a 256 c 266 g 296 t 34..others
ORIGIN

Alignment Scores:

Pred. No.: 3 83e-155 Length: 1201
Score: 1534.00 Matches: 300
Percent Similarity: 99.348 Conservative: 2
Best Local Similarity: 98.688 Mismatches: 1
Query Match: 85.278 Indels: 1
DB: 13 Gaps: 0

US-09-745-506-37 (1-350) x BX446370 (1-1201)

QY 48 GluGlualMetGluGluValleuGlnLysLysAlaAspleuIleleuserTYrHisPro 67

Db 67 GATGAAGTATGAGAGAGTCTCTGCAAAAGGACAGACTTCTCTCAACATCGG 126
QY 68 ProIlePheArgProMetLysArgIleThrTPasnThrTrpLysGluArgLeuValIle 87
Db 127 CCTATCTCCGACCCCTGTAAGGCGCATACCTGGAACACATGTGAAGAGCGCTGTGATC 186
QY 88 ATGAlaLeuGlnAsnArgValGlyIleTYrserProHisThrAlaIleTrpAlaIlePro 107
Db 187 CGGGCTCTGAGAACAGATGGGTATCTCTCTCCATACAGCTATATGCTGGCCC 246
QY 108 GlnGlyValAsnAsnTrpLeuAlaLysGlyLeuGlyAlaCysThrSerArgProIleHis 127
Db 247 CAGGGGCTCAACAACATGTTGGCTAAAGGCGTTGAGCTTGATACCTCAGGCCCATACAT 306
QY 128 ProSerLysAlaProAsnTYrProThrGluGlnAsnHisArgValGluPheAsnValAsn 147
Db 307 CTTTCCCAAGCTCCCACTACCTACAGAGGGAACCCAGATACAAATTCAACGTTAAC 366
QY 148 TYrThrGlnAspleuAspleuValMetSerAlaValIleAspGlyValIleSerVal 167
Db 367 TACACCAAGACCTGGACAAAGTATGCTGCHGTGAAGAAATTCAGCGTTTCTGTTC 426
QY 168 ThrSerPheSerAlaArgThrGlnAsnGlnGlnInThrArgIleAsnLeuAsnCysThr 187
Db 427 ACTCTTTTCTGCTGAGACTGCTGATGAGAAACAACAGGATTAATCGAATTGTACT 486
QY 188 GlnLysAlaLeuMetGlnValValAspPheLeuSerArgAsnLysGlnLeuTYrGlnLys 207
Db 487 CAGAAAGCTTTGATGAGGTGTGATTTTCTTCCCGAACAACAACTTTATCAAGAAG 546
QY 208 ThrGlnIleLeuSerLeuGlnLysProLeuLeuHisThrGlnMetGlyArgLeuGlnLys 227
Db 547 ACGAATATCTGCTGATGAGAGAGCTTTGCTTTCATATCTGAAATGGAGCGTTATGCG 606
QY 228 ThrLeuAspLysLeuSerValSerLeuAlaThrMetIleAspArgIleLysArgHisLeuLys 247
Db 607 ACACGTGAGAAATCTGTCTCCCGCAACATGATGATCAATAAAGACACCTAAAA 666
QY 248 LeuSerHisIleArgGlnLeuGlnValGlyArgThrLeuGlnLysGlnVal 267
Db 667 CATCTCATATTCGCTTACCTGAGGCTGGGGAGAACTTACAGTCAAGTCAAAGATC 726
QY 268 ValAlaLeuCysAlaGlySerGlySerSerValLeuGlnGlyValGluAlaAspleuTYr 287
Db 727 GTGGCCCTGTGTCTGTCTGTGTGGACAGGCTTCACAGGCTTACAGGTGACCTTAC 786
QY 288 LeuThrGlyGlnMetSerHisHisAspThrLeuAspAlaAlaSerGlnGlyIleAsnVal 307
Db 787 CTCACAGTGAGATGTCCCATCATGATCTTGGATGAGTGCCTCCCAAGAAATAAATGTC 846
QY 308 IleLeuCysGlnHisSerAsnThrGluValGlyPheLeuSerAspleuArgAspMetLeu 327
Db 847 ATCTCTGTGAAACAGCAACACTGAACGAGGCTTCTTCTGACCTTGACGATTAAGCTG 906
QY 328 AspSerHisLeuGlnLysLysIleAsnIleIle-leuSerGluThrAspArgAspProle 347
Db 907 GATTTCACCTTGAAGATAAAGAAATATATCTTATCTATCAGACATGAAAGGACCTCT 966
QY 347 uGlnValVal 350
Db 967 TCAGGTGTGTA 976

RESULT 5

LOCUS AL581997 1201 bp mRNA linear EST 01-JUN-2003
DEFINITION AL581997 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
Homo sapiens CDNA clone CSDBL003YGC24 5-PRIME, mRNA sequence.

ACCESSION AL581997 GI:31320228
VERSION AL581997.2
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

QY 180 rArgIlaaenleuancysThrGlnLysAlaLeuMetGlnValValaspheleuSer 200
 Db 756 CCGGATGAGCTGCTGAATGTACTCAGAAAGACTTGCAGAGCTTTCCTTCCCA 815
 QY 200 gAsnLysGlnLeuYrGlnLysThrGlnLeuSerLeuGlnLysProLeuLeuH1 220
 Db 816 GGACAGACAACTTATCAGAAAGCTGAATCTTTCATGAGAAAGCTTGTGCTGCA 875
 QY 220 sThrGlyMetGlyArgLeuCysThrLeuAspGlySerValSerLeuAlaThrMetLeu 240
 Db 876 TACGGAATGAGGAGGTGTGTCACACATGATGATCTGTCCCTCGCAATATGATAGA 935
 QY 240 pArgIleLysArgHisLeuLysLeuSerHisIleArgLeuAlaLeuGlyValGlyArg 260
 Db 936 GCGAATCAAAACACCTTAAAGCTGCGCATCTTCTGCTTACCTCTTGGAGTGGGAGAAC 995
 QY 260 rLeuGlnSerGlnValLysValAlaLeuCysAlaGlySerGlySerValLeuG1 280
 Db 996 ATTAGAGTCCCAAGTCAAAAGTGTGGCCCTGTGTGCTTGTGGGGGAGTGTTCRCA 1055
 QY 280 nGlyValGlnAlaAspLeuYrLeuThrGlyGlyMetSerHisIleAspThrLeuAsp 300
 Db 1056 AGGAGTGGAGCGCAGCTTCTACCTCACAGTGAATGTCCACCATGATGTTCGATGC 1115
 QY 300 aAlaSerGlnGlyIleAsnValIleLeuCysGlnHisSerAsnThrGlnArgGlyPhe 320
 Db 1116 TGTCTCCAAAGGATCAATGTGTCATCTTGTGAACACAGCAACCTGAAGAGCTTCT 1175
 QY 320 uSerAspLeuArgAspMetLeuAspSerHisLeuGlnLysIleAsnIleLeuLeu 340
 Db 1176 TTCTGAGCTTCAGAAAGTGTGGGTGTCTACTTGTAGATATAGATTAACTATTCCTGTC 1235
 QY 340 rGluThrAspArgAspProLeuGlnValVal 350
 Db 1236 TGAGACAGACAGGAGCCTCTCCGNGTGT 1266

RESULT 7

BX393871

LOCUS

DEFINITION BX393871 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens

ACCESSION CDNA clone CS0DC005YM21 5-PRIME, mRNA sequence.

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

EST. Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 L1 M.B., Gruber, C., Jessee, J. and Polayes, P.
 Full-length cDNA libraries and normalization
 Unpublished
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 1287.f for
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DC005AG110P1;cluster=1287.f. Contact :
 Feng Liang Email : fliang@lifetech.com URL : Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DC005AG110P1.
 Location/Qualifiers
 1..930
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DC005YM21"
 /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
 /clone_id="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)

FEATURES

source

primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT

248 a 231 c 214 g 237 t

ALIGNMENT Scores:

Pred. No.: 1,97e-138 Length: 930
 Score: 1378.00 Matches: 267
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 76.60% Indels: 0
 Gaps: 0

US-09-745-506-37 (1-350) x BX393871 (1-930)

QY 1 MetaspLeuYsaAlaLeuLeuSerSerLeuAsnAspPheAlaSerLeuSerPheAlaGlu 20
 Db 130 ATGATTTGAAAGGCTCTCTCTTCTCTCTGATACCTTGCATCCCTGCTGTGAG 189
 QY 21 SerTrpAspAsnValGlyLeuLeuValGluProSerProProHisThrValAsnThrLeu 40
 Db 190 AGTTGGGACAAATGTTGGATTACTGCTGGAAACCAAGCCACACACATATCTGTAATACACTC 249
 QY 41 PheLeuThrAsnAspLeuThrGlnGluValMetGlnGluValLeuGlnLysValAsp 60
 Db 250 TTCTGTGACCAATGACCTGCTGAGAGAGTGTGAGAGGCTGTGCAAAAGAGGACAGAC 309
 QY 61 LeuIleLeuSerTrpHisProProIlePheArgProMetLysArgIleThrTrpAsnThr 80
 Db 310 CTGATTCCTCTTACCATCCGCTTATCTCGGACCATGAAAGCCATATCTGGAACACA 369
 QY 81 TrpLysGluArgLeuValIleArgAlaLeuGlnLysAsnArgValGlyIleYrSerProHis 100
 Db 370 TGAAGAGAGCCCTGTGTATCCGGGCTGTGAGAAACAGAGTGTATCTACTCTCATAT 429
 QY 101 ThrAlaTrpAspAlaAlaProGlnGlyValAsnAsnTrpLeuAlaLysGlyLeuGlyAla 120
 Db 430 ACAGCTATGATGATGCTGCGCCCGCCAGGCGCTCAACAACATGTGTGCTAAAGGCTTGAGCT 489
 QY 121 CysTrpSerArgProIleHisProSerLysAlaProAsnTrpProThrGlnGlyAsnHis 140
 Db 490 TGTACTCTCAGGCCCATATCATCTTCCAAAGCTCCCAACTCCCTACAGAGGAACACAC 549
 QY 141 ArgValGluPheAsnValAsnTrpThrGlnAspLeuAspLysValMetSerAlaValLys 160
 Db 550 CGAGTAGAATTCACAGCTTACACTACCCCAAGACCTGGCAAAAGTCATGTGACAGTGAA 609
 QY 161 GlyIleAspGlyValSerValThrSerPheSerAlaArgThrGlyAsnGlnGlnThr 180
 Db 610 GGAATGACGGTGTCTGTCACTTCTTTCTGTCTGAGACTGTATAGAGAACCAACA 669
 QY 181 ArgIleAsnLeuAsnCysThrGlnLysAlaLeuMetGlnValValAspPheLeuSerArg 200
 Db 670 CGGATTTATTCGAATGTGACAGAAAGCTTGTGATGACGATGATGATTTCTTCCCG 729
 QY 201 AsnLysGlnLeuYrGlnLysThrGlnLeuSerLeuGlnLysProLeuLeuH1s 220
 Db 730 AACAAACAACTTATCAGAAAGCGAAATTTGTCTACGAGAGAGCCCTTCTCTTACAT 789
 QY 221 ThrGlyMetGlyArgLeuCysThrLeuAspGlySerValSerLeuAlaThrMetLeu 240
 Db 790 ACTGGAATGAGGAGGTGTGACACATGATGATGATGATGATGATGATGATGATGAT 849
 QY 241 ArgIleLysArgHisLeuLysLeuSerHisIleArgLeuAlaLeuGlyValGlyArg 260
 Db 850 CGAATTAACAAAGACCTTAAATATCTATTCATATTCCTTACGCCCTTGGGGTGGGAGAAC 909
 QY 261 LeuGlnSerGlnValLysVal 267
 Db 910 TTAGAGTCTCAAGTCAAGTGC 930

RESULT 8

BU172435
 LOCUS BU172435 857 bp mRNA linear EST 04-SEP-2002
 DEFINITION AGENCOURT 7970438 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6164871
 5', mRNA sequence.
 ACCESSION BU172435
 VERSION BU172435.1 GI:22686419
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 857)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC/DCPD/DPF
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1AM13522 row: k column: 16
 High quality sequence stop: 728.
 Location/Qualifiers
 1..857
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6164871"
 /rname="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC-72"
 /note="Organ: Skin; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 2 kb. Library constructed by Life
 Technologies."
 BASE COUNT 234 a 201 c 199 g 223 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 7.26e-137 Length: 857
 Score: 1363.00 Matches: 267
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 75.76% Indels: 0
 Gaps: 0
 US-09-745-506-37 (1-350) x BU172435 (1-857)
 QY 84 ATGLeuValIILeArgAlaLeuGluAsnArGValGlyILEtyrSerProH:sthrAlaTyr 103
 DB 3 CCGCTGGATCATCCGGGCTCTGGAGAACAGAGTCGTATCTCTCTCTACAGCCTAT 62
 QY 104 ASPAlaAlaProGlnGlyValAsnAsnTPleuAlaLeuGlyGlyLeuGlyAlaGlySer 123
 DB 63 GATGCTGCCGCCAGGCGCTCAACAACGTGGCTGAAAGGGCTTGAGCTGTGTACTCTCC 122
 QY 124 ATGProIleH1ProSerLysAlaProAsnTyrProThrGlnGlyAsnH1sarArGValGlu 143
 DB 123 AGGCCCATACATCTCTCCAAAGCTCCCACTACCCCTACAGAGAGAAACACGAGTAA 182
 QY 144 PheAsnValAsnTyrThrGlnAspLeuAspLysValMetSerAlaValLysGlyIleAsp 163
 DB 183 TTCACAGTAACTACACCCAGACCTCGAGCAAAAGTCATGTCTGCAAGGAATTCAC 242
 QY 164 G1ValSerValThrSerPheSerAlaArGThrGlyAsnGlnGlnGlnGlnGlnGlnGln 183
 DB 243 GGTGTTTCTGTCACCTCTTTCTGCTAGAGCTGGTAATGAGGACCAACACGAGATTAA 302
 QY 184 LeuAsnCysThrGlnLysAlaLeuMetGlnValValAspPheLeuSerArGAsnLysGln 203

DB 303 CTGAATTGATCAGAGAAGCGTTTGATGACAGGCTGTGATTTCTTCCGGAGCAACAA 362
 QY 204 LeuTyrGlnLysThrGlnIleLeuSerLeuGlnLysProLeuLeuH1sthrGlyMet 223
 DB 363 CTTTATCAGAAAGAGGAAATTCGTCTACTGAGAACGCTTGTCTTACTACGAAATG 422
 QY 224 G1ArgLeuCysThrLeuAspGluSerValSerLeuAlaThrMetIleAspArgIleLys 243
 DB 423 GGAGGTTATGACACATGATGATCTGTCTCCGCAACATGATGATGCAATPAAA 482
 QY 244 ARGH1sLeuLysLeuSerH1sLeuArgLeuAlaLeuGlyValGlyArgThrLeuGluSer 263
 DB 483 AGACACTAAACATCTCATCTGATTCCTTACGCTTGGGCTGGGAGAACCTTAGAGTCT 542
 QY 264 G1nValLysValAlaAlaLeuCysAlaGlySerGlySerSerValLeuGlnGlyValGlu 283
 DB 543 CAAGTCAAAATCCGTGGCCCTGTGCTGGTGTGGAGAGCAGCTTCTGCAGGCTGTAG 602
 QY 284 AlaAspLeuTyrThrGlnGlyMetSerH1sAspThrLeuAspAlaAlaSerGln 303
 DB 603 GCTGACCTTTTACCTCACAGGTGAGATGTCCATCATGATCTTGGATGCTCTCCAA 662
 QY 304 G1YIleAsnValIleLeuCysGlnH1sSerArThrGlnArgGlyPheLeuSerAspLeu 323
 DB 663 GGAATTAATGTCATCTCTGTGAAACACAGCAACACTGACGAGGCTTCTTGTGACCTT 722
 QY 324 ArgAspMetLeuAspSerH1sLeuGluAsnLysIleAsnIleLeuSerGlnThrAsp 343
 DB 723 CGAGATATGCTGATCTCTGAGAAATGATTAATATTCATATCAGAGACTGAC 782
 QY 344 ArgAspProLeuGlnValVal 350
 DB 783 AGGAGCCCTTCAGGTGTA 803
 RESULT 9
 BU859307 929 bp mRNA linear EST 16-OCT-2002
 LOCUS AGENCOURT_10422791 NIH_MGC_109 Homo sapiens cDNA clone
 DEFINITION IMAGE:6650148 5', mRNA sequence.
 ACCESSION BU859307
 VERSION BU859307.1 GI:24044299
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 929)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1CM2894 row: o column: 12
 High quality sequence stop: 754.
 Location/Qualifiers
 1..929
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6650148"
 /rname="teratocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC-109"
 /note="Organ: ovary; Vector: pORF7; Site_1: EcoRI; Site_2:
 XhoI; cDNA made by oligo-dT priming. Directionally cloned

|||||
Db 325 CTCATCTCTCTACCATCGGCTATCTTCGACCCATGAAGGCATTAACCTGGAAACA 384
Qy 81 TTPlySGluArgLeuValIleArgAlaLeuGluAsnArgValGlyIleTyrSerProHis 100
Db 385 TGGAGAGAGCGCTGGTGGATCCGGGCTCTGGAGAACAGACTCGGTATCTCTCCCTCAT 444
Qy 101 ThrAlaTyrAspAlaAlaProGlnGlyValAsnAspTyrLeuAlaIleGlyLeuGlyAla 120
Db 445 ACAGCCTATGATCTGCGCCCGCCAGGGGCTACACACAGGTTGGCTAAAGGGCTGGAGCT 504
Qy 121 CysThrSerArgProIleHisProSerTyrAlaProAsnTyrProThrGlnGlyAsnHis 140
Db 505 TGTACCTCCAGGCCCATACATCTCTCCAAAGCTCCCACTACCTACAGAGGGAACACAC 564
Qy 141 ArgValGluPheAsnValAsnTyrThrGlnAspLeuAspIleValMetSerAlaValLys 160
Db 565 CGAGTAGAATTCACCTTACCTACACCCAGACCTGGACAAAGTCAATGTCTGCAGTGA 624
Qy 161 GlyIleAspGlyValSerValThrSerPheSerAlaArgThrGlnGlyAsnGlnGlnThr 180
Db 625 GGAATTGACGGCTTTCTCTGCTACTTCTTTTCTGCTAGGACTGGTGGTGGAGAACACA 684
Qy 181 ArgIleAsnLeuAsnCysThrGlnLysAlaLeuMetGlnValValAspPheLeuSerArg 200
Db 685 CGGATTAATCTGAATGTACTCAGAAAGCTTGTATGAGAGTGGATTTTCTTCCCGG 744
Qy 201 AsnIleGlnLeuTyrGlnLysThrGlnIleLeuSerLeuGlnLysProLeuLeuLeuHis 220
Db 745 AACAAACAACTTATATCAGAAAGAGGAATTCCTGCTACCTGGAGAGCCTTGGCTTCAT 804
Qy 221 ThrGlyMetGlyArgLeuCysThrLeuAspGluSerValSerLeuAlaThrMetIleAsp 240
Db 805 ACAGGATGGAGCGGTTATGACACACTGGATGATCTCTCTCCCTGCAACACATGATTTGAT 864
Qy 241 ArgIleLysArgHisLeuLysLeuSerHisIleArgLeuAlaLeuGlnValGlyArgThr 260
Db 865 CGAATGAAAAGACACCTAANAACTATCTCATATTCGCTTACCCCTTGGGGGGGGAGAAC 924
Qy 261 -leuGluSerGlnValLysValAlaIle-leuCysAlaGlySerGlySer 277
Db 925 CTTAGAGTCTCAAGTCAAAAGTCTGGCCCTGTGTCTGGTCTGGAAAGAAC 977

RESULT 11
BQ229243 836 bp mRNA linear EST 02-MAY-2002
LOCUS BQ229243
DEFINITION AGENCOURT_7510797 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6055296
5', mRNA sequence.
ACCESSION BQ229243
VERSION BQ229243.1 GI:20410643
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 836)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapds-remail.nih.gov
Tissue Procurement: ATCC/DC/DMP
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: L1AM1315 row: f column: 01
High quality sequence stop: 677.
Location/Qualifiers
1..836
/organism="Homo sapiens"

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6055296"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_72"
/note="Organ: skin; Vector: pCMV-Sport6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally. primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 218 a 205 c 200 g 211 t 2 others
ORIGIN
Alignment Scores:
Pred. No.: 1..3e-136 Length: 836
Score: 1360.50 Matches: 268
Percent Similarity: 97.45% Conservative: 0
Best Local Similarity: 97.45% Mismatches: 5
Query Match: 75.63% Indels: 2
DB: 13 Gaps: 1
US-09-745-506-37 (1-350) x BQ229243 (1-836)
Qy 3 LeuLysAlaLeuLeuSerSerLeuAsnAspPheAlaSerLeuSerPheAlaGluSerTyr 22
Db 2 TTGAGGCT 61
Qy 23 AspAsnValGlyLeuLeuValGluProSerProHisThrValAsnThrLeuPheLeu 42
Db 62 GACAACTTGTGATTAAGTGGGAGAACCAAGCCACACATCACTGTAATTAACATCTTCTG 121
Qy 43 ThrAsnAspLeuThrGlnGlnValMetGlnGlnValLeuGlnLysLysAlaAspLeuIle 62
Db 122 ACCAATACCTGACTGAGAGATGATGAGAGAGTGTCTGCAAAAGAGCGACCTGATT 181
Qy 63 LeuSerTyrHisProProIlePheArgProMetLysArgIleThrPAsnThrTyrLys 82
Db 182 CTCTCTACATCCGCGCTATCTCTCCAGACCATGAAGGCATTAACCTGGAACATGTGAG 241
Qy 83 GluArgLeuValIleArgAlaLeuGluAsnArgValGlyIleTyrSerProHisThrAla 102
Db 242 GAGGCGCTGTGATCCGGGCTCTGGAGAACAGACTCGGTATCTCTCTCAATACAGCC 301
Qy 103 TyrAspAlaAlaProGlnGlyValAsnAspTyrLeuAlaIleGlyLeuGlyAlaCysThr 122
Db 302 TATGATCTCTGCGCCCGCCAGGCGCTCAACATCGTTGGCTTAAGGCTTGGACTTTGACC 361
Qy 123 SerArgProIleHisProSerTyrAlaProAsnTyrProThrGlnGlyAsnHisArgVal 142
Db 362 TCCAGGCCCATACATCTCTCCAAAGCTCCCACTACCTTCAGAGGGAACACCGAGTA 421
Qy 143 GluPheAsnValAsnTyrThrGlnAspLeuAspIleValMetSerAlaValLysGlyIle 162
Db 422 GAATTCACCTTAATCACTACACCCAGACCTGACAAAGTCAATGTCTGAGTGAAGGATTT 481
Qy 163 AspGlyValSerValThrSerPheSerAlaArgThrGlnGlnGlnGlnGlnGlnGlnGln 182
Db 482 GAGCGTCTTCTGCTCACTTCTTTTCTGCTAGAGCTGTAATGAGAACAAACCGGATT 541
Qy 183 AsnLeuAsnCysThrGlnLysAlaLeuMetGlnValValAspPheLeuSerArgAsnLys 202
Db 542 AACTTGAATTTGCTACTCAGAAAGGCTTGTATGAGTGTGATATTTCTTCCCGGAACAAA 601
Qy 203 GlnLeuTyrGlnLysThrGlnIleLeuSerLeuGlnLysProLeuLeuHisThrGly 222
Db 602 CAATTTTATCAGAAAGAGGAATTCCTGCTACGAGAGAGCTTGGCTTTCATACATACGGA 661
Qy 223 MetGlyArgLeuCysThrLeuAspGluSerValSerLeuAlaThrMetIleAspArgIle 242
Db 662 ATGGAGAGGTTATGACACATCGATGATCTGTCTCCCTGGCAACCATGATGATGATGATA 721
Qy 243 LysArgHisLeuLysLeuSerHisIleArgLeuAlaLeuGlnValGlyArgThrLeuGln 262
|||||

Db 722 AAAAGACCTAAATATCTCATATTCCTTACCCCTTGGGGGGGAGAACCTTGA 781

Qy 262 userglnvallyvalalaleucysala---GlySergly 275
|||||
782 GCTCAAGTCAAGTGGNGGCCCTGTGCTGCTGTTNTGGG 824

RESULT 12
LOCUS AL582089/c 1201 bp mRNA linear EST 01-JUN-2003
DEFINITION Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
Homo sapiens cDNA clone CS0DD003YG24 3-PRIME, mRNA sequence.
AL582089
ACCESSION AL582089.2 GI:31320317
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1201)
AUTHORS L.W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT On Feb 16, 2001 this sequence version replaced gi:12949732.
Contact: Genoscope - Centre National de Sequencage
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by life technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1287.f for
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DD003BD12NP1&cluster=1287.f. Contact :
Feng Liang Email : fliang@life.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DD003BD12NP1.
location/qualifiers
1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DD003YG24"
/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/cell_line="RAMOS CELL LINE"
/clone_id="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 309 a 259 c 260 g 321 t 52 others

ORIGIN

Alignment Scores:

Pred. No.:	9.75e-135	Length:	1201
Score:	1345.50	Matches:	280
Percent Similarity:	91.23%	Conservative:	1
Best local Similarity:	90.91%	Mismatches:	26
Query Match:	74.79%	Indels:	4
DB:	9	Gaps:	0

US-09-745-506-37 (1-350) x AL582089 (1-1201)

Qy 44 AsnaSplLeuthrGluGluValMetGluGluValLeuGlnLysLysAlaSplLeuLeu 63
|||||
Db 1135 AATGACCTAACTAGAGAGAGTAKKGRGGGG-KTGAAAGAAAGAAACCTTAATTC- 1078
|||||

Qy 64 SerTyRHsPProPlePhe-ArgPrometLysArgLietHrTPasHrTTPlysgl 83
|||||
Db 1077 TCCTACCAATCCGCTTTCTCCCAACCATRAAGGCAATACCTGGAAACATGAGRGA 1018
|||||

Qy 83 uArgLeuValIleatYtAlaLeuGlnAsnArgValGlyIleTySerProHsThAlatY 103
|||||
Db 1017 GCGCTGTGATCCGGGTYCTGGAGACAGAGTCCGATCTCTCTCCATACACCTTA 958
|||||

Qy 103 rAspAlaIaProGlnGlyValAsnaSntPleuAlaIleGlyLeuGlnValIaCysThrse 123
|||||
Db 957 TTGATSTGSCCCAGGCGGCGCAACACTGTGTGCTTAAGGGCTTGAGCTKT-ACCCTC 899
|||||

Qy 123 rArgProIleHsPProSerLysAlaProAsnTyProThrGluGlnLysHsArgValgl 143
|||||
Db 898 CAGGCCCATCATCTCTTCCAAAGCTCCCACTACCTACCTACAGAGGAAACCGAGTGA 839
|||||

Qy 143 uPheAsnValAsnTyThrGlnAspLeuAspLysValIleMetSerAlaValysGlyIleAs 163
|||||
Db 838 ATTCACAGTAACTACACCCCAAGACCTGGACAAAGTATKTCTCAGTGAAGGAATGA 779
|||||

Qy 163 pGlyValSerValHrSerPheSerAlaArgThrGlyAsnGlnGluGlnHrArgIleAs 183
|||||
Db 778 CGGTGTTCTGCTACTCTTTTCTGCTAGAGCTGTATGAGAACAAACCGATTA 719
|||||

Qy 183 nLeuAsnCysThrGlnLysAlaLeuMetGlnValValAspPheLeuSerArgAsnLysgl 203
|||||
Db 718 TCTGAATTTGCTCAGAAAGCTTTGATGCAGCTGTGATTTCTTCCCGGAACAACA 659
|||||

Qy 203 nLeuTyGlnLysThrGluIleLeuSerLeuGlnLysProLeuLeuHsIsthrglyme 223
|||||
Db 658 ACTTTATCAGAAAGCGAAATTTCTGCTACGTGAGAGACCTTGTCTACATAGTGAAT 599
|||||

Qy 223 tGlyArgLeuGlnLysThrLeuAspGlnSerValSerLeuAlaHrMetIleAspArgIlely 243
|||||
Db 598 GGGACGGTTATGACACACAGTGAATGTCTCCCTGGCAACATGATGATGCAATAA 539
|||||

Qy 243 sArgHsIleuLysLeuSerHsIleArgLeuAlaLeuGlnValGlyArgThrLeuGlnse 263
|||||
Db 538 AAGACACCTTAATCATATCTCATATTCGCTTACCTTGGGGTGGGGAACCTTAGACTC 479
|||||

Qy 263 rGlnValLysValAlaLeuGlnLysGlySerglySerSerValLeuGlnGlyValgl 283
|||||
Db 478 TCAAGTCAAGACGCGGGCCCTGTGCTGTGCTGGAGACAGCTTGTGCGAGGGGTGTA 419
|||||

Qy 283 uAlaAspLeuTyThrLeuThrGlnGluMetSerHsIsthAspHrLeuAspAlaIaSerGl 303
|||||
Db 418 GCGTACCTTTACCTCAGAGGTGAGATGTCCCATATATATCTTGATGCTGCTTCCCA 359
|||||

Qy 303 nGlyIleAsnValIleLeuGlnLysGlnHsIsthAsnThrGlnArgIleLeuSerAspLe 323
|||||
Db 358 AGGATTAATATGCATCTCTGTGACACAGACAGACACTGACGASCTCTCTCTGACCT 299
|||||

Qy 323 uArgAspMetLeuAspSerHsIsthLeuGlnLysLysIleAsnIleIleLeuSerGlnThAs 343
|||||
Db 298 TCGAGATATGCTGGANNCTCACTGTGAGATTAAGATTAATATCTATCAGAGACTGA 239
|||||

Qy 343 pArgAspProLeuGlnValVal 350
|||||
Db 238 CAGGAGCCTCKTCAAGTGTGA 217
|||||

RESULT 13
LOCUS AL563496/c 1201 bp mRNA linear EST 31-MAY-2003
DEFINITION AL563496 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
cDNA clone CS0DD006YH18 3-PRIME, mRNA sequence.
AL563496
ACCESSION AL563496.2 GI:31287490
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1201)
AUTHORS L.W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT On Feb 15, 2001 this sequence version replaced gi:12912946.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France

|||||
Db 882 ACTGGATGGAGCGGTGGACACCTGATGATCTGCTCCCTGGCAATGAATAGAG 941
Qy 241 ArgilleysArgHisLeuLeuSerHisIleArgLeuAlaLeuGlyValArgThr 260
Db 942 CGAATGAAAACACACTTAAGCTGTGCATCTTCGCTTACCTTGTGAGTGGGAGAAC 1001
Qy 261 LeuGluSerGlnValValAlaLeuGlyAlaGlySerGlySerValLeuGln 280
Db 1002 TTAGAGTCCCAAGTCAAAAGTTGGCCCTGTGTCTGTGGGGCAGGTGTTCTACAA 1061
Qy 281 GtValGluAlaLeuLeuThrGlyGluMetSerHisHisSphrLeuAsp 299
Db 1062 GGAGTGGAGCGCAGCTTACCTACAGAGCAGTGTCTTGGAACAGAGACGTGAA 1118

RESULT 15
AL521920 957 bp mRNA linear EST 22-MAY-2003
LOCUS AL521920 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens
DEFINITION CDNA clone CS0DB003YPI4 5-PRIME, mRNA sequence.
ACCESSION AL521920
VERSION AL521920.2 GI:31040194
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 957)
Li,W.B., Gruber,C., Jesssee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 13, 2001 this sequence version replaced gi:12785413.
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1287.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DB003DH07OP1&cluster=1287.f Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DB003DH07CPI.
FEATURES
source
Location/Qualifiers
1..957
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DB003YPI4"
/issue_type="NEUROBLASTOMA COT 10-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 250 a 239 c 217 g 250 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 1,36e-131 Length: 957
Score: 1315.00 Matches: 261
Percent Similarity: 98.86 Conservative: 0
Best Local Similarity: 98.86 Mismatches: 3
Query Match: 73.10e Indels: 2
Gaps: 0

US-09-745-506-37 (1-350) x AL521920 (1-957)

Qy 1 MetAspLeuValAlaLeuLeuSerLeuAsnAspPheAlaSerLeuSarPheAlaGlu 20
Db 166 ATGAGATTGAAGGCTCTCTTCTTCTTCTTGAATGACTTTGCAATCCCTCTGCTTGGCTGAG 225

Qy 21 SerTPAspAsnValGlyLeuLeuValGluProSerProHisThrValAsnThrLeu 40
Db 226 AGTTGGACAAATGTGGATTACTGTGGAAACCAAGCCACACATGTAATATACACTC 285
Qy 41 PheLeuThrAsnAspLeuThrGluValMetGluGluValLeuGlnLysValAsp 60
Db 286 TTCCTGACCAATACCTGACTGAGAGATGATGAGAGGTGTGCAAAAAGAAAGGCAGAC 345
Qy 61 LeuIleLeuSerThrHisProPheArgProMetLysArgIleThrPAsnThr 80
Db 346 CTCATCTCTCCACACATCCGCGCTATCTCCAGACCATGAGGCAATAACTGGAAACACA 405
Qy 81 TPPLysGluArgLeuValIleArgAlaLeuGluAsnArgValGlyIleYrSerProHis 100
Db 406 TGGAGAGAGCGCTGGATCCGGGCTCTGGAACAGAGTCCGTAATCTACTCTCCAT 465
Qy 101 ThrAlaThrAspAlaProGlnGlyValAsnAspThrLeuValLysGlyLeuGlyAla 120
Db 466 ACAGCTATGATCTCGCGCCCGAGGCGTCACAACTGTTGGCTAAAGGGCTTGGAGCT 525
Qy 121 CysThrSerArgProIleHisProSerLysAlaProAsnThrProThrGluGlyAsnHis 140
Db 526 TGTACCTCCAGGCCCATACATCTTCCAAAGCTCCCAACTACCTACAGAGGAAACAC 585
Qy 141 ArgValGluPheAsnValAsnThrThrGlnAspLeuAspLysValMetSerAlaValLys 160
Db 586 CGAGTAGAATTCAACGTTACTACACCAAGACCTGGACAAAGTCAATGCTGCAGTGA 645
Qy 161 GlyIleAspGlyValSerValThrSerPheSerAlaArgThrGlnGlnGluGlnThr 180
Db 646 GGAATTGACGCTGTCTCTCACTTCTTTTCTGTAGAGACTGTAATGGAACAACA 705
Qy 181 ArgIleAsnLeuAsnCysThrGlnLysAlaLeuMetGlnValValAspPheLeuSerArg 200
Db 706 CGGATTAACTGAATGTACTACAGAGGCTTGTGATGAGGTGATGATTTCTTCCCGG 765
Qy 201 AsnLysGlnLeuThrGlnLysThrGluIleLeuSerLeuGluLysProLeuLeuHis 220
Db 766 AACCAACAACTTTATCAGAAAGCAGAAATCTGTCACTGGAGAAAGCCTTGTGCTCAT 825
Qy 221 ThrGlyMetGlyArgLeuCysThrIleAspGluSerValSerLeuAlaThrMetIleAsp 240
Db 826 ACTGGATGGAGCGTTATGACACATGATGAAATCTGTCTCCCTGGCAACCATGATGAT 885
Qy 241 ArgIleLysArgHisLeuLysLeuSerHisIleArgLeuAlaLeuGlyValGlyArgThr 260
Db 886 CGAAT-AAAAGACACCTAAACATATCTCATATTCGCTTACCCCTTGGGGAGGAGAACT 944
Qy 261 LeuGluSerGln 264
Db 945 TA-GAGTCTCAA 955

Search completed: August 23, 2003, 19:55:01
Job time : 1911 secs

THIS PAGE BLANK (USPTO)